CLAIMS

What is claimed is:

1. A method for generating three-dimensional models from uncalibrated views comprising acts of:

- forming a three-dimensional model of at least a portion of a scene viewed from an uncalibrated image capturing device location by:
  - receiving images from uncalibrated views of the uncalibrated image capturing device location;
  - extracting features from the images from uncalibrated views of the uncalibrated image capturing device location;
  - computing correspondence between features from images from the uncalibrated views captured from the uncalibrated image capturing device location;
  - forming a three-dimensional structure modeling the at least a portion of a scene viewed from the uncalibrated image capturing device location;
- iteratively performing the act of forming a three-dimensional model viewed from an uncalibrated image capturing device location, for a subset of the uncalibrated image capturing device locations available;
- stitching together the three-dimensional models viewed from the subset of the uncalibrated image capturing device locations by:
  - finding spatially local persistent feature groupings from the uncalibrated views captured at an uncalibrated image capturing device location;
  - iteratively performing the act of finding spatially local persistent feature groupings for a subset of the uncalibrated image capturing device locations available;
  - computing correspondence between sets of feature groupings from two uncalibrated image capturing device locations for a subset of

pair-wise combinations of uncalibrated image capturing device locations;

- searching for best matches, whereby multiple three-dimensional models from a subset of uncalibrated image capturing device locations are thus "stitched" together to form an overall three-dimensional model of at least a portion of a scene; and
- outputting the overall three-dimensional model of at least a portion of a scene.

2. A method for generating three-dimensional models from uncalibrated views as set forth in claim 1, wherein in the act of receiving images from uncalibrated views, the images are obtained using at least one uncalibrated image capturing device selected from a group consisting of a still camera, a video camera, a Magnetic Resonance Imaging (MRI) recording mechanism, an ultrasound recording apparatus, an external computer connected to the system, an internet connection, an internet web camera, a direct satellite link, a video cassette recorder (VCR), a digital versatile disc (DVD) player, and imaging recording media used to gather snapshots of a desired portion of a scene at multiple uncalibrated views, the imaging recording media selected from a group consisting of a Compact Disk (CD), a Digital Versatile Disk / Digital Video Disk (DVD), a floppy disk, a magnetic tape, a removable hard drive, a printed picture, a scanned document, a faxed document, a digital camera, a video cassette, a Magnetic Resonance Imaging (MRI) recording media, an ultrasound recording media, and a solid-state recording media.

3. A method for generating three-dimensional models from uncalibrated views as set forth in claim 2, wherein the images from uncalibrated views are generated from a group consisting of: images generated by a single uncalibrated image capturing device viewing at least a portion of a scene

at multiple pan and tilt settings; images captured with multiple uncalibrated image capturing devices viewing at least a portion of a scene; and images generated by multiple uncalibrated image capturing devices viewing at least a portion of a scene at multiple pan and tilt settings.

4. A method for generating three-dimensional models from uncalibrated views as set forth in claim 3, wherein a portion of a scene comprises at least one object, and wherein the images from uncalibrated views are formed from a group consisting of: images containing overlapping views of a portion of a scene, images containing partially overlapping views of a portion of a scene, images containing slightly overlapping views of a portion of a scene, and images containing non-overlapping views of a portion of a scene.

5. A method for generating three-dimensional models from uncalibrated views as set forth in claim 4, further comprising a act of identifying and eliminating unpaired features prior to computing correspondence between features and computing correspondence between sets of feature groupings.

6. A method for generating three-dimensional models from uncalibrated views as set forth in claim 5, wherein in the act of extracting features from the images, the features include at least one of: corner features, high entropy points, local edge features, and contour features.

7. A method for generating three-dimensional models from uncalibrated views as set forth in claim 6, wherein the correspondence between features and the correspondence between sets of feature groupings are computed by using a technique selected from a group consisting of: probabilistic matching, correlation measure, chi-square statistical measure, and dot product of feature vectors.

8. A method for generating three-dimensional models from uncalibrated views as set forth in claim 7, wherein in the act of forming a three-dimensional structure from the uncalibrated image capturing device location, the three-dimensional structure is formed by a "structure from motion" algorithm, where motion from the uncalibrated image capturing device is computed from the correspondence established from the images of uncalibrated views captured at different pan-tilt settings of the uncalibrated image capturing device location, and where the "structure from motion" algorithm simulates a three-dimensional structure modeling at least a portion of a scene from the motion of the uncalibrated image capturing device.

9. A method for generating three-dimensional models from uncalibrated views as set forth in claim 8, wherein the correspondence between features and the correspondence between sets of feature groupings are computed using a probabilistic matching method, where the probabilistic matching method computes probabilities of match between features by using prior information representing a portion of a scene, and where the probabilities of match between features correspond to *a posteriori* probabilities.

10. A method for generating three-dimensional models from uncalibrated views as set forth in claim 9, wherein in the act of identifying and eliminating unpaired features, two unpaired features are identified by computing and plotting an *a posteriori* probability relating both features, wherein the *a posteriori* probability relating both features has a flat profile when the two features are unpaired.

11. A method for generating three-dimensional models from uncalibrated views as set forth in claim 10, wherein the *a posteriori* probabilities are

used to form a correspondence matrix, where a one-to-one correspondence between feature groupings from two uncalibrated image capturing device locations is established by maximizing the correspondence matrix.

12. A method for generating three-dimensional models from uncalibrated views as set forth in claim 11, wherein a Sinkhorn normalization process is used to form the correspondence matrix.

13. A method for generating three-dimensional models from uncalibrated views as set forth in claim 12, wherein in the act of searching for the best matches, a Ransac robust estimation algorithm is used to find peaks on the correspondence matrix, wherein the peaks on the correspondence matrix indicate where the three-dimensional models from the uncalibrated image capturing device locations are to be stitched together.

14. A method for generating three-dimensional models from uncalibrated views as set forth in claim 13, wherein the *a posteriori* probability, $P(H_i \mid X)$, is defined by

$$P(H_i \mid X) = \sum_{k=1}^{K} P(H_i \mid X, \xi_{X,\mu_k}) P(\xi_{X,\mu_j} \mid X = X_n)$$

where

$X = [X_1, \ldots, X_N]$ and $Y = [Y_1, \ldots, Y_M]$ denote two sets of features extracted from two images from uncalibrated views; wherein a total of N features were extracted from an uncalibrated view, and a total of M features were extracted from another uncalibrated view;

$\mu = \mu_1, \ldots, \mu_K$ represent a set of features extracted from prior information, wherein a total of K features were extracted from prior information of the portion of the scene;

H_i denotes a hypothesis that a feature Y_i matches the set of features X, wherein i is a variable index with an integer value between 1 and M, and wherein i denotes an i^{th} feature within the set of features Y;

$\xi_{X\mu_j}$ denotes an event that $\{X$ matches $\mu_j\}$, wherein j is a variable index with an integer value between 1 and K, and wherein j denotes a j^{th} feature within the set of prior features μ;

$P(\xi_{X,\mu_j} \mid X = X_n)$ denotes a probability of the set of features X matching a prior information feature μ_j given that the set of features X consists of a feature X_n, wherein n is a variable index with an integer value chosen between 1 and N, and wherein n denotes a n^{th} feature within the set of features X;

$P(H_i \mid X, \xi_{X,\mu_j})$ denotes a probability of a feature Y_i matching the set of features X given an event $\xi_{X\mu_j}$ occurred (i.e. the set of features X matches the prior information feature μ_j); wherein i denotes the i^{th} feature within the set of features Y and j denotes the j^{th} feature within the set of prior features μ;

and wherein the probability of the set of features X matching the prior information feature μ_j, $P(\xi_{X,\mu_j} \mid X = X_n)$, and the probability of a feature Y_i matching the set of features X given an event $\xi_{X\mu_j}$ occurred, $P(H_i \mid X, \xi_{X,\mu_j})$, are defined by

$$P(\xi_{X,\mu_j} \mid X = X_n) = \frac{1}{\sum_{k=1}^{K} < X_n, \mu_k >} < X_n, \mu_j >$$

and

$$P(H_i \mid X, \xi_{X,\mu_j}) = \frac{1}{\sum_{k=1}^{K} E < Y_i, \mu_k >} < Y_i, \mu_j >$$

where

E denotes a probabilistic expectation measure, and $< \cdot >$ denotes an inner product, where the inner product represents a measure of similarity.

15. A method for generating three-dimensional models from uncalibrated views as set forth in claim 14, wherein in the act of outputting the overall three-dimensional model, the outputting device is selected from a group consisting of at least one of: a computer monitor, a video camera connected to a computer, and a computer readable media used to display the overall three-dimensional model of a portion of a scene, the computer readable media selected from a group consisting of an imaging Compact Disk (CD), a Digital Versatile Disk/ Digital Video Disk (DVD), a floppy disk, a removable hard drive, a video cassette, and a solid-state recording media.

16. A method for generating three-dimensional models from uncalibrated views as set forth in claim 1, wherein the images from uncalibrated views are generated from a group consisting of: images generated by a single uncalibrated image capturing device viewing at least a portion of a scene at multiple pan and tilt settings; images captured with multiple uncalibrated image capturing devices viewing at least a portion of a scene; and images generated by multiple uncalibrated image capturing devices viewing at least a portion of a scene at multiple pan and tilt settings.

17. A method for generating three-dimensional models from uncalibrated views as set forth in claim 1, wherein a portion of a scene comprises at

least one object, and wherein the images from uncalibrated views are formed from a group consisting of: images containing overlapping views of a portion of a scene, images containing partially overlapping views of a portion of a scene, images containing slightly overlapping views of a portion of a scene, and images containing non-overlapping views of a portion of a scene.

18. A method for generating three-dimensional models from uncalibrated views as set forth in claim 1, further comprising a act of identifying and eliminating unpaired features prior to computing correspondence between features and computing correspondence between sets of feature groupings.

19. A method for generating three-dimensional models from uncalibrated views as set forth in claim 1, wherein in the act of extracting features from the images, the features include at least one of: corner features, high entropy points, local edge features, and contour features.

20. A method for generating three-dimensional models from uncalibrated views as set forth in claim 1, wherein in the act of forming a three-dimensional structure from the uncalibrated image capturing device location, the three-dimensional structure is formed by a "structure from motion" algorithm, where motion from the uncalibrated image capturing device is computed from the correspondence established from the images of uncalibrated views captured at different pan-tilt settings of the uncalibrated image capturing device location, and where the "structure from motion" algorithm simulates a three-dimensional structure modeling at least a portion of a scene from the motion of the uncalibrated image capturing device.

21. A method for generating three-dimensional models from uncalibrated views as set forth in claim 1, wherein the correspondence between features and the correspondence between sets of feature groupings are computed by using a technique selected from a group consisting of: probabilistic matching, correlation measure, chi-square statistical measure, and dot product of feature vectors.

22. A method for generating three-dimensional models from uncalibrated views as set forth in claim 1, wherein the correspondence between features and the correspondence between sets of feature groupings are computed using a probabilistic matching method, where the probabilistic matching method computes probabilities of match between features by using prior information representing a portion of a scene, and where the probabilities of match between features correspond to *a posteriori* probabilities.

23. A method for generating three-dimensional models from uncalibrated views as set forth in claim 22, wherein in the act of identifying and eliminating unpaired features, two unpaired features are identified by computing and plotting an *a posteriori* probability relating both features, wherein the *a posteriori* probability relating both features has a flat profile when the two features are unpaired.

24. A method for generating three-dimensional models from uncalibrated views as set forth in claim 23, wherein the *a posteriori* probabilities are used to form a correspondence matrix, where a one-to-one correspondence between feature groupings from two uncalibrated image capturing device locations is established by maximizing the correspondence matrix.

25. A method for generating three-dimensional models from uncalibrated views as set forth in claim 22, wherein the *a posteriori* probability, $P(H_i \mid X)$, is defined by

$$P(H_i \mid X) = \sum_{k=1}^{K} P(H_i \mid X, \xi_{X,\mu_k}) P(\xi_{X,\mu_j} \mid X = X_n)$$

where

$X = [X_1, \ldots, X_N]$ and $Y = [Y_1, \ldots, Y_M]$ denote two sets of features extracted from two images from uncalibrated views; wherein a total of N features were extracted from an uncalibrated view, and a total of M features were extracted from another uncalibrated view;

$\mu = \mu_1, \ldots, \mu_K$ represent a set of features extracted from prior information, wherein a total of K features were extracted from prior information of the portion of the scene;

H_i denotes a hypothesis that a feature Y_i matches the set of features X, wherein i is a variable index with an integer value between 1 and M, and wherein i denotes an i^{th} feature within the set of features Y;

$\xi_{X\mu_j}$ denotes an event that { X matches μ_j }, wherein j is a variable index with an integer value between 1 and K, and wherein j denotes a j^{th} feature within the set of prior features μ;

$P(\xi_{X,\mu_j} \mid X = X_n)$ denotes a probability of the set of features X matching a prior information feature μ_j given that the set of features X consists of a feature X_n, wherein n is a variable index with an integer value chosen between 1 and N, and wherein n denotes a n^{th} feature within the set of features X;

$P(H_i \mid X, \xi_{X,\mu_j})$ denotes a probability of a feature Y_i matching the set of features X given an event $\xi_{X\mu_j}$ occurred (i.e. the set of features

X matches the prior information feature μ_j); wherein i denotes the i^{th} feature within the set of features Y and j denotes the j^{th} feature within the set of prior features μ ;

and wherein the probability of the set of features X matching the prior information feature μ_j , $P(\xi_{X,\mu_j} \mid X = X_n)$, and the probability of a feature Y_i matching the set of features X given an event $\xi_{X\mu_j}$ occurred, $P(H_i \mid X, \xi_{X,\mu_j})$, are defined by

$$P(\xi_{X,\mu_j} \mid X = X_n) = \frac{1}{\sum_{k=1}^{K} < X_n, \mu_k >} < X_n, \mu_j >$$

and

$$P(H_i \mid X, \xi_{X,\mu_j}) = \frac{1}{\sum_{k=1}^{K} E < Y_i, \mu_k >} < Y_i, \mu_j >$$

where

E denotes a probabilistic expectation measure, and $< \ . \ >$ denotes an inner product, where the inner product represents a measure of similarity.

26. A method for generating three-dimensional models from uncalibrated views as set forth in claim 1, wherein a Sinkhorn normalization process is used to form the correspondence matrix.

27. A method for generating three-dimensional models from uncalibrated views as set forth in claim 1, wherein in the act of searching for the best matches, a Ransac robust estimation algorithm is used to find peaks on the correspondence matrix, wherein the peaks on the correspondence matrix

indicate where the three-dimensional models from the uncalibrated image capturing device locations are to be stitched together.

28. A method for generating three-dimensional models from uncalibrated views as set forth in claim 1, wherein in the act of outputting the overall three-dimensional model, the outputting device is selected from a group consisting of at least one of: a computer monitor, a video camera connected to a computer, and a computer readable media used to display the overall three-dimensional model of a portion of a scene, the computer readable media selected from a group consisting of an imaging Compact Disk (CD), a Digital Versatile Disk/ Digital Video Disk (DVD), a floppy disk, a removable hard drive, a video cassette, and a solid-state recording media.

29. A system for generating three-dimensional models from uncalibrated views, the system comprising:

a computer system including a processor, a memory coupled with the processor, an input coupled with the processor for receiving images from uncalibrated views of an uncalibrated image capturing device location, the computer system further comprising means, residing in its processor and memory for:

- forming a three-dimensional model of at least a portion of a scene viewed from an uncalibrated image capturing device location by:
  - extracting features from the images from uncalibrated views of the uncalibrated image capturing device location;
  - computing correspondence between features from images from the uncalibrated views captured from the uncalibrated image capturing device location;

- forming a three-dimensional structure modeling the at least a portion of a scene viewed from the uncalibrated image capturing device location;
- iteratively performing the means for forming a three-dimensional model viewed from an uncalibrated image capturing device location, for a subset of the uncalibrated image capturing device locations available;
- stitching together the three-dimensional models viewed from the subset of the uncalibrated image capturing device locations by:
  - finding spatially local persistent feature groupings from the uncalibrated views captured at an uncalibrated image capturing device location;
  - iteratively performing the means for finding spatially local persistent feature groupings for a subset of the uncalibrated image capturing device locations available;
  - computing correspondence between sets of feature groupings from two uncalibrated image capturing device locations for a subset of pair-wise combinations of uncalibrated image capturing device locations;
  - searching for best matches, whereby multiple three-dimensional models from a subset of uncalibrated image capturing device locations are thus "stitched" together to form an overall three-dimensional model of at least a portion of a scene; and
- outputting the overall three-dimensional model of at least a portion of a scene.

30. A system for generating three-dimensional models from uncalibrated views as set forth in claim 29, wherein in the means for receiving images from uncalibrated views, the images are obtained using at least one uncalibrated image capturing device selected from a group consisting of a

still camera, a video camera, a Magnetic Resonance Imaging (MRI) recording mechanism, an ultrasound recording apparatus, an external computer connected to the system, an internet connection, an internet web camera, a direct satellite link, a video cassette recorder (VCR), a digital versatile disc (DVD) player, and imaging recording media used to gather snapshots of a desired portion of a scene at multiple uncalibrated views, the imaging recording media selected from a group consisting of a Compact Disk (CD), a Digital Versatile Disk / Digital Video Disk (DVD), a floppy disk, a magnetic tape, a removable hard drive, a printed picture, a scanned document, a faxed document, a digital camera, a video cassette, a Magnetic Resonance Imaging (MRI) recording media, an ultrasound recording media, and a solid-state recording media.

31. A system for generating three-dimensional models from uncalibrated views as set forth in claim 30, wherein the images from uncalibrated views are generated from a group consisting of: images generated by a single uncalibrated image capturing device viewing at least a portion of a scene at multiple pan and tilt settings; images captured with multiple uncalibrated image capturing devices viewing at least a portion of a scene; and images generated by multiple uncalibrated image capturing devices viewing at least a portion of a scene at multiple pan and tilt settings.

32. A system for generating three-dimensional models from uncalibrated views as set forth in claim 31, wherein a portion of a scene comprises at least one object, and wherein the images from uncalibrated views are formed from a group consisting of: images containing overlapping views of a portion of a scene, images containing partially overlapping views of a portion of a scene, images containing slightly overlapping views of a portion of a scene, and images containing non-overlapping views of a portion of a scene.

33. A system for generating three-dimensional models from uncalibrated views as set forth in claim 32, further comprising a means for identifying and eliminating unpaired features prior to computing correspondence between features and computing correspondence between sets of feature groupings.

34. A system for generating three-dimensional models from uncalibrated views as set forth in claim 33, wherein in the means for extracting features from the images, the features include at least one of: corner features, high entropy points, local edge features, and contour features.

35. A system for generating three-dimensional models from uncalibrated views as set forth in claim 34, wherein the correspondence between features and the correspondence between sets of feature groupings are computed by using a technique selected from a group consisting of: probabilistic matching, correlation measure, chi-square statistical measure, and dot product of feature vectors.

36. A system for generating three-dimensional models from uncalibrated views as set forth in claim 35, wherein in the means for forming a three-dimensional structure from the uncalibrated image capturing device location, the three-dimensional structure is formed by a "structure from motion" algorithm, where motion from the uncalibrated image capturing device is computed from the correspondence established from the images of uncalibrated views captured at different pan-tilt settings of the uncalibrated image capturing device location, and where the "structure from motion" algorithm simulates a three-dimensional structure modeling at least a portion of a scene from the motion of the uncalibrated image capturing device.

37. A system for generating three-dimensional models from uncalibrated views as set forth in claim 36, wherein the correspondence between features and the correspondence between sets of feature groupings are computed using a probabilistic matching method, where the probabilistic matching method computes probabilities of match between features by using prior information representing a portion of a scene, and where the probabilities of match between features correspond to *a posteriori* probabilities.

38. A system for generating three-dimensional models from uncalibrated views as set forth in claim 37, wherein in the means for identifying and eliminating unpaired features, two unpaired features are identified by computing and plotting an *a posteriori* probability relating both features, wherein the *a posteriori* probability relating both features has a flat profile when the two features are unpaired.

39. A system for generating three-dimensional models from uncalibrated views as set forth in claim 38, wherein the *a posteriori* probabilities are used to form a correspondence matrix, where a one-to-one correspondence between feature groupings from two uncalibrated image capturing device locations is established by maximizing the correspondence matrix.

40. A system for generating three-dimensional models from uncalibrated views as set forth in claim 39, wherein a Sinkhorn normalization process is used to form the correspondence matrix.

41. A system for generating three-dimensional models from uncalibrated views as set forth in claim 40, wherein in the means for searching for the best matches, a Ransac robust estimation algorithm is used to find peaks

on the correspondence matrix, wherein the peaks on the correspondence matrix indicate where the three-dimensional models from the uncalibrated image capturing device locations are to be stitched together.

42. A system for generating three-dimensional models from uncalibrated views as set forth in claim 41, wherein the *a posteriori* probability, $P(H_i \mid X)$, is defined by

$$P(H_i \mid X) = \sum_{k=1}^{K} P(H_i \mid X, \xi_{X,\mu_k}) P(\xi_{X,\mu_j} \mid X = X_n)$$

where

$X = [X_1, \ldots, X_N]$ and $Y = [Y_1, \ldots, Y_M]$ denote two sets of features extracted from two images from uncalibrated views; wherein a total of N features were extracted from an uncalibrated view, and a total of M features were extracted from another uncalibrated view;

$\mu = \mu_1, \ldots, \mu_K$ represent a set of features extracted from prior information, wherein a total of K features were extracted from prior information of the portion of the scene;

H_i denotes a hypothesis that a feature Y_i matches the set of features X, wherein i is a variable index with an integer value between 1 and M, and wherein i denotes an i^{th} feature within the set of features Y;

$\xi_{X\mu_j}$ denotes an event that { X matches μ_j }, wherein j is a variable index with an integer value between 1 and K, and wherein j denotes a j^{th} feature within the set of prior features μ;

$P(\xi_{X,\mu_j} \mid X = X_n)$ denotes a probability of the set of features X matching a prior information feature μ_j given that the set of features X consists of a feature X_n, wherein n is a variable index with an

integer value chosen between *1* and *N*, and wherein *n* denotes a n^{th} feature within the set of features *X*;

$P(H_i \mid X, \xi_{X,\mu_j})$ denotes a probability of a feature Y_i matching the set of features *X* given an event $\xi_{X\mu_j}$ occurred (i.e. the set of features X matches the prior information feature μ_j); wherein *i* denotes the i^{th} feature within the set of features *Y* and *j* denotes the j^{th} feature within the set of prior features μ;

and wherein the probability of the set of features *X* matching the prior information feature μ_j, $P(\xi_{X,\mu_j} \mid X = X_n)$, and the probability of a feature Y_i matching the set of features *X* given an event $\xi_{X\mu_j}$ occurred, $P(H_i \mid X, \xi_{X,\mu_j})$, are defined by

$$P(\xi_{X,\mu_j} \mid X = X_n) = \frac{1}{\sum_{k=1}^{K} < X_n, \mu_k >} < X_n, \mu_j >$$

and

$$P(H_i \mid X, \xi_{X,\mu_j}) = \frac{1}{\sum_{k=1}^{K} E < Y_i, \mu_k >} < Y_i, \mu_j >$$

where

E denotes a probabilistic expectation measure, and **< . >** denotes an inner product, where the inner product represents a measure of similarity.

43. A system for generating three-dimensional models from uncalibrated views as set forth in claim 42, wherein in the means for outputting the overall three-dimensional model, the outputting device is selected from a group consisting of at least one of: a computer monitor, a video camera

connected to a computer, and a computer readable media used to display the overall three-dimensional model of a portion of a scene, the computer readable media selected from a group consisting of an imaging Compact Disk (CD), a Digital Versatile Disk/ Digital Video Disk (DVD), a floppy disk, a removable hard drive, a video cassette, and a solid-state recording media.

44. A system for generating three-dimensional models from uncalibrated views as set forth in claim 29, wherein the images from uncalibrated views are generated from a group consisting of: images generated by a single uncalibrated image capturing device viewing at least a portion of a scene at multiple pan and tilt settings; images captured with multiple uncalibrated image capturing devices viewing at least a portion of a scene; and images generated by multiple uncalibrated image capturing devices viewing at least a portion of a scene at multiple pan and tilt settings.

45. A system for generating three-dimensional models from uncalibrated views as set forth in claim 29, wherein a portion of a scene comprises at least one object, and wherein the images from uncalibrated views are formed from a group consisting of: images containing overlapping views of a portion of a scene, images containing partially overlapping views of a portion of a scene, images containing slightly overlapping views of a portion of a scene, and images containing non-overlapping views of a portion of a scene.

46. A system for generating three-dimensional models from uncalibrated views as set forth in claim 29, further comprising a means for identifying and eliminating unpaired features prior to computing correspondence between features and computing correspondence between sets of feature groupings.

47. A system for generating three-dimensional models from uncalibrated views as set forth in claim 29, wherein in the means for extracting features from the images, the features include at least one of: corner features, high entropy points, local edge features, and contour features.

48. A system for generating three-dimensional models from uncalibrated views as set forth in claim 29, wherein in the means for forming a three-dimensional structure from the uncalibrated image capturing device location, the three-dimensional structure is formed by a "structure from motion" algorithm, where motion from the uncalibrated image capturing device is computed from the correspondence established from the images of uncalibrated views captured at different pan-tilt settings of the uncalibrated image capturing device location, and where the "structure from motion" algorithm simulates a three-dimensional structure modeling at least a portion of a scene from the motion of the uncalibrated image capturing device.

49. A system for generating three-dimensional models from uncalibrated views as set forth in claim 29, wherein the correspondence between features and the correspondence between sets of feature groupings are computed by using a technique selected from a group consisting of: probabilistic matching, correlation measure, chi-square statistical measure, and dot product of feature vectors.

50. A system for generating three-dimensional models from uncalibrated views as set forth in claim 29, wherein the correspondence between features and the correspondence between sets of feature groupings are computed using a probabilistic matching method, where the probabilistic matching method computes probabilities of match between features by

using prior information representing a portion of a scene, and where the probabilities of match between features correspond to *a posteriori* probabilities.

51. A system for generating three-dimensional models from uncalibrated views as set forth in claim 50, wherein in the means for identifying and eliminating unpaired features, two unpaired features are identified by computing and plotting an *a posteriori* probability relating both features, wherein the *a posteriori* probability relating both features has a flat profile when the two features are unpaired.

52. A system for generating three-dimensional models from uncalibrated views as set forth in claim 51, wherein the *a posteriori* probabilities are used to form a correspondence matrix, where a one-to-one correspondence between feature groupings from two uncalibrated image capturing device locations is established by maximizing the correspondence matrix.

53. A system for generating three-dimensional models from uncalibrated views as set forth in claim 50, wherein the *a posteriori* probability, $P(H_i \mid X)$, is defined by

$$P(H_i \mid X) = \sum_{k=1}^{K} P(H_i \mid X, \xi_{X,\mu_k}) P(\xi_{X,\mu_j} \mid X = X_n)$$

where

$X = [X_1, \ldots, X_N]$ and $Y = [Y_1, \ldots, Y_M]$ denote two sets of features extracted from two images from uncalibrated views; wherein a total of N features were extracted from an uncalibrated view, and a total of M features were extracted from another uncalibrated view;

$\mu = \mu_1, ..., \mu_K$ represent a set of features extracted from prior information, wherein a total of K features were extracted from prior information of the portion of the scene;

H_i denotes a hypothesis that a feature Y_i matches the set of features X, wherein i is a variable index with an integer value between 1 and M, and wherein i denotes an i^{th} feature within the set of features Y;

$\xi_{X\mu_j}$ denotes an event that { X matches μ_j }, wherein j is a variable index with an integer value between 1 and K, and wherein j denotes a j^{th} feature within the set of prior features μ ;

$P(\xi_{X,\mu_j} \mid X = X_n)$ denotes a probability of the set of features X matching a prior information feature μ_j given that the set of features X consists of a feature X_n, wherein n is a variable index with an integer value chosen between 1 and N, and wherein n denotes a n^{th} feature within the set of features X;

$P(H_i \mid X, \xi_{X,\mu_j})$ denotes a probability of a feature Y_i matching the set of features X given an event $\xi_{X\mu_j}$ occurred (i.e. the set of features X matches the prior information feature μ_j); wherein i denotes the i^{th} feature within the set of features Y and j denotes the j^{th} feature within the set of prior features μ ;

and wherein the probability of the set of features X matching the prior information feature μ_j , $P(\xi_{X,\mu_j} \mid X = X_n)$, and the probability of a feature Y_i matching the set of features X given an event $\xi_{X\mu_j}$ occurred, $P(H_i \mid X, \xi_{X,\mu_j})$, are defined by

$$P(\xi_{X,\mu_j} \mid X = X_n) = \frac{1}{\sum_{k=1}^{K} < X_n, \mu_k >} < X_n, \mu_j >$$

and

$$P(H_i \mid X, \xi_{X,\mu_j}) = \frac{1}{\sum_{k=1}^{K} E < Y_i, \mu_k >} < Y_i, \mu_j >$$

where

E denotes a probabilistic expectation measure, and $< \cdot >$ denotes an inner product, where the inner product represents a measure of similarity.

54. A system for generating three-dimensional models from uncalibrated views as set forth in claim 29, wherein a Sinkhorn normalization process is used to form the correspondence matrix.

55. A system for generating three-dimensional models from uncalibrated views as set forth in claim 29, wherein in the means for searching for the best matches, a Ransac robust estimation algorithm is used to find peaks on the correspondence matrix, wherein the peaks on the correspondence matrix indicate where the three-dimensional models from the uncalibrated image capturing device locations are to be stitched together.

56. A system for generating three-dimensional models from uncalibrated views as set forth in claim 29, wherein in the means for outputting the overall three-dimensional model, the outputting device is selected from a group consisting of at least one of: a computer monitor, a video camera connected to a computer, and a computer readable media used to display the overall three-dimensional model of a portion of a scene, the computer readable media selected from a group consisting of an imaging Compact

Disk (CD), a Digital Versatile Disk/ Digital Video Disk (DVD), a floppy disk, a removable hard drive, a video cassette, and a solid-state recording media.

57. A computer program product for generating three-dimensional models from uncalibrated views, the computer program product comprising means, stored on a computer readable medium for:

- forming a three-dimensional model of at least a portion of a scene viewed from an uncalibrated image capturing device location by:
  - receiving images from uncalibrated views of the uncalibrated image capturing device location;
  - extracting features from the images from uncalibrated views of the uncalibrated image capturing device location;
  - computing correspondence between features from images from the uncalibrated views captured from the uncalibrated image capturing device location;
  - forming a three-dimensional structure modeling the at least a portion of a scene viewed from the uncalibrated image capturing device location;
- iteratively performing the means for forming a three-dimensional model viewed from an uncalibrated image capturing device location, for a subset of the uncalibrated image capturing device locations available;
- stitching together the three-dimensional models viewed from the subset of the uncalibrated image capturing device locations by:
  - finding spatially local persistent feature groupings from the uncalibrated views captured at an uncalibrated image capturing device location;

- iteratively performing the means for finding spatially local persistent feature groupings for a subset of the uncalibrated image capturing device locations available;
- computing correspondence between sets of feature groupings from two uncalibrated image capturing device locations for a subset of pair-wise combinations of uncalibrated image capturing device locations;
- searching for best matches, whereby multiple three-dimensional models from a subset of uncalibrated image capturing device locations are thus "stitched" together to form an overall three-dimensional model of at least a portion of a scene; and
- outputting the overall three-dimensional model of at least a portion of a scene.

58. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 57, wherein in the means for receiving images from uncalibrated views, the images are obtained using at least one uncalibrated image capturing device selected from a group consisting of a still camera, a video camera, a Magnetic Resonance Imaging (MRI) recording mechanism, an ultrasound recording apparatus, an external computer connected to the system, an internet connection, an internet web camera, a direct satellite link, a video cassette recorder (VCR), a digital versatile disc (DVD) player, and imaging recording media used to gather snapshots of a desired portion of a scene at multiple uncalibrated views, the imaging recording media selected from a group consisting of a Compact Disk (CD), a Digital Versatile Disk / Digital Video Disk (DVD), a floppy disk, a magnetic tape, a removable hard drive, a printed picture, a scanned document, a faxed document, a digital camera, a video cassette, a Magnetic Resonance Imaging (MRI) recording media, an ultrasound recording media, and a solid-state recording media.

59. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 58, wherein the images from uncalibrated views are generated from a group consisting of: images generated by a single uncalibrated image capturing device viewing at least a portion of a scene at multiple pan and tilt settings; images captured with multiple uncalibrated image capturing devices viewing at least a portion of a scene; and images generated by multiple uncalibrated image capturing devices viewing at least a portion of a scene at multiple pan and tilt settings.

60. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 59, wherein a portion of a scene comprises at least one object, and wherein the images from uncalibrated views are formed from a group consisting of: images containing overlapping views of a portion of a scene, images containing partially overlapping views of a portion of a scene, images containing slightly overlapping views of a portion of a scene, and images containing non-overlapping views of a portion of a scene.

61. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 60, further comprising a means for identifying and eliminating unpaired features prior to computing correspondence between features and computing correspondence between sets of feature groupings.

62. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 61, wherein in the means for extracting features from the images, the features include at least one of:

corner features, high entropy points, local edge features, and contour features.

63. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 62, wherein the correspondence between features and the correspondence between sets of feature groupings are computed by using a technique selected from a group consisting of: probabilistic matching, correlation measure, chi-square statistical measure, and dot product of feature vectors.

64. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 63, wherein in the means for forming a three-dimensional structure from the uncalibrated image capturing device location, the three-dimensional structure is formed by a "structure from motion" algorithm, where motion from the uncalibrated image capturing device is computed from the correspondence established from the images of uncalibrated views captured at different pan-tilt settings of the uncalibrated image capturing device location, and where the "structure from motion" algorithm simulates a three-dimensional structure modeling at least a portion of a scene from the motion of the uncalibrated image capturing device.

65. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 64, wherein the correspondence between features and the correspondence between sets of feature groupings are computed using a probabilistic matching method, where the probabilistic matching method computes probabilities of match between features by using prior information representing a portion of a scene, and where the probabilities of match between features correspond to *a posteriori* probabilities.

66. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 65, wherein in the means for identifying and eliminating unpaired features, two unpaired features are identified by computing and plotting an *a posteriori* probability relating both features, wherein the *a posteriori* probability relating both features has a flat profile when the two features are unpaired.

67. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 66, wherein the *a posteriori* probabilities are used to form a correspondence matrix, where a one-to-one correspondence between feature groupings from two uncalibrated image capturing device locations is established by maximizing the correspondence matrix.

68. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 67, wherein a Sinkhorn normalization process is used to form the correspondence matrix.

69. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 68, wherein in the means for searching for the best matches, a Ransac robust estimation algorithm is used to find peaks on the correspondence matrix, wherein the peaks on the correspondence matrix indicate where the three-dimensional models from the uncalibrated image capturing device locations are to be stitched together.

70. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 69, wherein the *a posteriori* probability, $P(H_i \mid X)$, is defined by

$$P(H_i \mid X) = \sum_{k=1}^{K} P(H_i \mid X, \xi_{X,\mu_k}) P(\xi_{X,\mu_j} \mid X = X_n)$$

where

$X = [X_1, \ldots, X_N]$ and $Y = [Y_1, \ldots, Y_M]$ denote two sets of features extracted from two images from uncalibrated views; wherein a total of N features were extracted from an uncalibrated view, and a total of M features were extracted from another uncalibrated view;

$\mu = \mu_1, \ldots, \mu_K$ represent a set of features extracted from prior information, wherein a total of K features were extracted from prior information of the portion of the scene;

H_i denotes a hypothesis that a feature Y_i matches the set of features X, wherein i is a variable index with an integer value between 1 and M, and wherein i denotes an i^{th} feature within the set of features Y;

$\xi_{X\mu_j}$ denotes an event that $\{X$ matches $\mu_j\}$, wherein j is a variable index with an integer value between 1 and K, and wherein j denotes a j^{th} feature within the set of prior features μ;

$P(\xi_{X,\mu_j} \mid X = X_n)$ denotes a probability of the set of features X matching a prior information feature μ_j given that the set of features X consists of a feature X_n, wherein n is a variable index with an integer value chosen between 1 and N, and wherein n denotes a n^{th} feature within the set of features X;

$P(H_i \mid X, \xi_{X,\mu_j})$ denotes a probability of a feature Y_i matching the set of features X given an event $\xi_{X\mu_j}$ occurred (i.e. the set of features

X matches the prior information feature μ_j); wherein i denotes the i^{th} feature within the set of features Y and j denotes the j^{th} feature within the set of prior features μ ;

and wherein the probability of the set of features X matching the prior information feature μ_j , $P(\xi_{X,\mu_j} \mid X = X_n)$, and the probability of a feature Y_i matching the set of features X given an event $\xi_{X\mu_j}$ occurred, $P(H_i \mid X, \xi_{X,\mu_j})$, are defined by

$$P(\xi_{X,\mu_j} \mid X = X_n) = \frac{1}{\sum_{k=1}^{K} < X_n, \mu_k >} < X_n, \mu_j >$$

and

$$P(H_i \mid X, \xi_{X,\mu_j}) = \frac{1}{\sum_{k=1}^{K} E < Y_i, \mu_k >} < Y_i, \mu_j >$$

where

E denotes a probabilistic expectation measure, and $< \ . \ >$ denotes an inner product, where the inner product represents a measure of similarity.

71. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 70, wherein in the means for receiving images from uncalibrated views, the images are obtained using at least one uncalibrated image capturing device selected from a group consisting of a still camera, a video camera, a Magnetic Resonance Imaging (MRI) recording mechanism, an ultrasound recording apparatus, an external computer connected to the system, an internet connection, an internet web camera, a direct satellite link, a video cassette recorder (VCR), a digital versatile disc (DVD) player, and imaging recording

media used to gather snapshots of a desired portion of a scene at multiple uncalibrated views, the imaging recording media selected from a group consisting of a Compact Disk (CD), a Digital Versatile Disk / Digital Video Disk (DVD), a floppy disk, a magnetic tape, a removable hard drive, a printed picture, a scanned document, a faxed document, a digital camera, a video cassette, a Magnetic Resonance Imaging (MRI) recording media, an ultrasound recording media, and a solid-state recording media.

72. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 57, wherein the images from uncalibrated views are generated from a group consisting of: images generated by a single uncalibrated image capturing device viewing at least a portion of a scene at multiple pan and tilt settings; images captured with multiple uncalibrated image capturing devices viewing at least a portion of a scene; and images generated by multiple uncalibrated image capturing devices viewing at least a portion of a scene at multiple pan and tilt settings.

73. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 57, wherein a portion of a scene comprises at least one object, and wherein the images from uncalibrated views are formed from a group consisting of: images containing overlapping views of a portion of a scene, images containing partially overlapping views of a portion of a scene, images containing slightly overlapping views of a portion of a scene, and images containing non-overlapping views of a portion of a scene.

74. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 57, further comprising a means for identifying and eliminating unpaired features prior to

computing correspondence between features and computing correspondence between sets of feature groupings.

75. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 57, wherein in the means for extracting features from the images, the features include at least one of: corner features, high entropy points, local edge features, and contour features.

76. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 57, wherein in the means for forming a three-dimensional structure from the uncalibrated image capturing device location, the three-dimensional structure is formed by a "structure from motion" algorithm, where motion from the uncalibrated image capturing device is computed from the correspondence established from the images of uncalibrated views captured at different pan-tilt settings of the uncalibrated image capturing device location, and where the "structure from motion" algorithm simulates a three-dimensional structure modeling at least a portion of a scene from the motion of the uncalibrated image capturing device.

77. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 57, wherein the correspondence between features and the correspondence between sets of feature groupings are computed by using a technique selected from a group consisting of: probabilistic matching, correlation measure, chi-square statistical measure, and dot product of feature vectors.

78. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 57, wherein the

correspondence between features and the correspondence between sets of feature groupings are computed using a probabilistic matching method, where the probabilistic matching method computes probabilities of match between features by using prior information representing a portion of a scene, and where the probabilities of match between features correspond to *a posteriori* probabilities.

79. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 78, wherein in the means for identifying and eliminating unpaired features, two unpaired features are identified by computing and plotting an *a posteriori* probability relating both features, wherein the *a posteriori* probability relating both features has a flat profile when the two features are unpaired.

80. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 79, wherein the *a posteriori* probabilities are used to form a correspondence matrix, where a one-to-one correspondence between feature groupings from two uncalibrated image capturing device locations is established by maximizing the correspondence matrix.

81. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 78, wherein the *a posteriori* probability, $P(H_i \mid X)$, is defined by

$$P(H_i \mid X) = \sum_{k=1}^{K} P(H_i \mid X, \xi_{X,\mu_k}) P(\xi_{X,\mu_j} \mid X = X_n)$$

where

$X = [X_1, \ldots, X_N]$ and $Y = [Y_1, \ldots, Y_M]$ denote two sets of features extracted from two images from uncalibrated views; wherein a total of

N features were extracted from an uncalibrated view, and a total of *M* features were extracted from another uncalibrated view;

$\mu = \mu_1, \ldots, \mu_K$ represent a set of features extracted from prior information, wherein a total of *K* features were extracted from prior information of the portion of the scene;

H_i denotes a hypothesis that a feature Y_i matches the set of features *X*, wherein *i* is a variable index with an integer value between *1* and *M*, and wherein *i* denotes an i^{th} feature within the set of features *Y*;

$\xi_{X\mu_j}$ denotes an event that { *X* matches μ_j }, wherein *j* is a variable index with an integer value between *1* and *K*, and wherein *j* denotes a j^{th} feature within the set of prior features μ ;

$P(\xi_{X,\mu_j} \mid X = X_n)$ denotes a probability of the set of features *X* matching a prior information feature μ_j given that the set of features *X* consists of a feature X_n, wherein *n* is a variable index with an integer value chosen between *1* and *N*, and wherein *n* denotes a n^{th} feature within the set of features *X*;

$P(H_i \mid X, \xi_{X,\mu_j})$ denotes a probability of a feature Y_i matching the set of features *X* given an event $\xi_{X\mu_j}$ occurred (i.e. the set of features X matches the prior information feature μ_j); wherein *i* denotes the i^{th} feature within the set of features *Y* and *j* denotes the j^{th} feature within the set of prior features μ ;

and wherein the probability of the set of features *X* matching the prior information feature μ_j , $P(\xi_{X,\mu_j} \mid X = X_n)$, and the probability of a

feature Y_i matching the set of features X given an event $\xi_{X\mu_j}$ occurred, $P(H_i \mid X, \xi_{X,\mu_j})$, are defined by

$$P(\xi_{X,\mu_j} \mid X = X_n) = \frac{1}{\sum_{k=1}^{K} < X_n, \mu_k >} < X_n, \mu_j >$$

and

$$P(H_i \mid X, \xi_{X,\mu_j}) = \frac{1}{\sum_{k=1}^{K} E < Y_i, \mu_k >} < Y_i, \mu_j >$$

where

E denotes a probabilistic expectation measure, and **< . >** denotes an inner product, where the inner product represents a measure of similarity.

82. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 57, wherein a Sinkhorn normalization process is used to form the correspondence matrix.

83. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 57, wherein in the means for searching for the best matches, a Ransac robust estimation algorithm is used to find peaks on the correspondence matrix, wherein the peaks on the correspondence matrix indicate where the three-dimensional models from the uncalibrated image capturing device locations are to be stitched together.

84. A computer program product for generating three-dimensional models from uncalibrated views as set forth in claim 57, wherein in the means for outputting the overall three-dimensional model, the outputting device is

selected from a group consisting of at least one of: a computer monitor, a video camera connected to a computer, and a computer readable media used to display the overall three-dimensional model of a portion of a scene, the computer readable media selected from a group consisting of an imaging Compact Disk (CD), a Digital Versatile Disk/ Digital Video Disk (DVD), a floppy disk, a removable hard drive, a video cassette, and a solid-state recording media.